

## The role of prosodic focus in the reanalysis of garden path sentences: Depth of semantic processing impedes the revision of an erroneous local analysis

**June Choe**, Department of Linguistics, University of Pennsylvania, USA, [yjchoe@sas.upenn.edu](mailto:yjchoe@sas.upenn.edu)

**Masaya Yoshida**, Department of Linguistics, Northwestern University, USA, [m-yoshida@northwestern.edu](mailto:m-yoshida@northwestern.edu)

**Jennifer Cole**, Department of Linguistics, Northwestern University, USA, [jennifer.cole1@northwestern.edu](mailto:jennifer.cole1@northwestern.edu)

Research on post-repair representations of garden path sentences has found that readers systematically arrive at misinterpretations even after displaying evidence of reanalysis (Christianson et al., 2001; Ferreira et al., 2001). These comprehension errors have been attributed to the semantic interpretation associated with the incorrect parse persisting past disambiguation, but less is known about the mechanism driving this phenomenon (Sturt, 2007; Slattery et al., 2013). A speeded auditory comprehension experiment examined the depth of semantic processing as an independent influence on the strength of semantic persistence, drawing on known effects of pitch accent on the processing of focus-related semantic meaning (Fraundorf et al., 2010). Participants heard garden path sentences with early/late-closure ambiguity (e.g., *While Anna dressed the baby stopped crying*) with a sharply rising pitch accent on either the unambiguous adjunct subject or the ambiguously transitive adjunct verb, followed by a comprehension question that probed whether the incorrect late-closure analysis persisted. Since the pitch accent is often a strong cue for semantic focus when it occurs in prosodically marked phrase-medial positions, we reasoned that a deeper semantic processing would be facilitated for the late-closure analysis only when the verb receives a pitch accent. Findings indicate that a pitch accent on the verb significantly decreased accuracy without a corresponding increase on response time, suggesting that a deeper semantic processing of the erroneous parse can strengthen its resistance to revision without necessarily interfering with the process of structure-building.



## 1. Introduction

Despite the complexity of information in speech, utterances are typically comprehended rapidly, accurately, and effortlessly in conversations. This remarkable ability has been attributed in part to the incrementality of the human sentence processing mechanism, which allows the representation of an utterance to be built up as it unfolds in real time (Kimball, 1973; Abney, 1989; Crocker, 1996; Phillips, 1996; Townsend & Bever, 2001; Phillips & Lau, 2004; Phillips & Lewis, 2013). But for the same reason, the human parser faces difficulty processing local ambiguities that can only be resolved with information that comes later in the sentence. These ambiguities characterize the so-called *garden path* sentences, where the incorrect local parse is also the preferred parse that is initially pursued (thus leading the parser “down the garden path”). This has disruptive consequences for on-line structure building and often requires special attention via the process of reanalysis for the correct structure to be recovered (Frazier & Rayner, 1982, 1987). Interestingly, when the interpretation from the incorrect parse is locally coherent and semantically consistent with the rest of the sentence, this parsing error may sometimes go unnoticed (Christianson et al., 2001; Ferreira et al., 2001; Tabor & Hutchins, 2004; Tabor et al., 2004). In other words, comprehenders experience the illusion of having understood the sentence that was given to them, though the true underlying structure generates a different interpretation. Several models of this phenomenon have been proposed in the literature, where the debate has primarily centered on the question of whether the parser fails at recovering the true underlying structure or at discarding the initial semantic interpretation from the incorrect local parse (Ferreira et al., 2004; Tabor et al., 2004; Sturt, 2007; Slattery et al., 2013; Fujita, 2021). In this paper, we explore this question by testing whether the depth of semantic processing within the incorrect local parse independently drives the misinterpretation of garden path sentences, above and beyond the contribution from lexically driven local coherence effects that have been the focus of previous works. In a speeded auditory comprehension experiment, we manipulate the depth of semantic processing by drawing on a well-known relationship between pitch accent and semantic processing, where a rising pitch accent is a strong cue for focus in phrase-medial positions, but not when it occurs phrase-finally, in the default position of phrasal stress (Büring, 2016; Wagner, 2020).

## 2. Background

### 2.1 Misinterpretations in garden path sentences

The processing of temporary, local syntactic ambiguity in garden path sentences has been well studied in the psycholinguistics literature as a window into the internal mechanisms of the human parser (e.g., Bever, 1970; Frazier & Fodor, 1978; Frazier, 1979). Consider the sentence in (1), where the complementizer *while* signals the start of an adjunct clause. Because the presence of an adjunct clause entails the presence of the matrix clause that dominates it, the beginning of

the sentence signals the parser to expect a sentence which is minimally composed of two clauses – an adjunct clause followed by the matrix clause (Kazanina et al., 2007). However, up through the second NP, headed by *baby* (bolded region), it is ambiguous whether the boundary between the two clauses comes early after *dressed* or late after *cute*, as both (1a) and (1b) are possible continuations of the sentence fragment. Here, if the parser commits to the late-closure analysis of the adjunct clause VP but the actual structure of the sentence is one where that VP closes early, this presents a serious challenge for comprehension.

- (1) While Anna dressed **the baby that was small and cute** ...
- a. ... spat up on the bed. (early closure)
  - b. ... the neighbor visited. (late closure)

Fortunately, the parser is not deterministic such that it simply breaks down when it is led down the garden path (Marcus, 1980; Berwick & Weinberg, 1984). Instead, it has the capacity to recover from an initial misparse and build the correct underlying structure, though often with great effort (Fodor & Frazier, 1980; Marcus et al., 1983; Pritchett, 1992; Schneider & Phillips, 2001; Sturt et al., 2001). In reading studies, comprehenders have been reported to systematically display patterned behavior while processing ambiguous regions of the sentence that reflect this process of recovery from being garden-pathed, also called *reanalysis*. For example, Frazier and Rayner (1982) examined the on-line behavioral responses to parsing difficulty in sentences like (2) in an eye-tracking-while-reading paradigm.

- (2) Since Jay always jogs a mile seems like a very short distance to him.

Frazier and Rayner (1982) found that the parser initially interprets *a mile* as the direct object of *jogs*, although the correct analysis requires the adjunct clause VP to be closed early at the verb, with the NP *a mile* being analyzed as the subject of the matrix clause instead. This violation of the parser's expectation is reflected in longer looking times at the disambiguating region *seems like*, followed by regressive eye movements to an earlier region of the sentence. This sequence of behaviors has been taken to be a sign of a successful reanalysis because readers do not appear to experience any further difficulty reading the rest of the sentence. Representationally, this reanalysis process appears to allow the parser to both discard the initially incorrect analysis from further consideration and also build the globally correct analysis. In sum, reanalysis was understood to be a mechanism that allows comprehenders to arrive at the correct interpretation of garden path sentences despite the initial difficulty processing them.

However, this all-or-nothing view of reanalysis has been challenged by studies that more closely examine the post-repair representations of garden path sentences (Gibson, 1991; Gibson & Pearlmutter, 1998, 2000; Ferreira, Christianson, & Hollingworth, 2001; Ferreira, Bailey, &

Ferraro, 2002; Ferreira, Lau, & Bailey, 2004). For example, Christianson et al. (2001) reports that after reading sentences like (3), comprehenders often failed to correctly answer simple questions like (4) even after reporting high confidence in their answers.

(3) While Anna dressed the baby played in the crib.

(4) Did Anna dress the baby?

If reanalysis indeed facilitated the parser's recovery from its initial mistake of attaching *the baby* as the direct object of *dressed*, then the post-repair representation of the sentence should no longer license the interpretation that Anna dressed the baby. Therefore, these comprehension errors provide evidence that the process of reanalysis as traditionally understood could sometimes fail at yielding the correct interpretation of the sentence. At the same time, however, comprehenders consistently answered questions like (5) correctly, suggesting that the failure of reanalysis is very particular and local.

(5) Did the baby play in the crib?

In light of this finding, early works by Ferreira and colleagues have made a broader claim about the relationship between sentence processing, the syntactic structure, and comprehension (e.g., Ferreira et al., 2001). In Ferreira et al., (2002), the syntactic structure is claimed to be fragile and not by itself a reliable source of sentence meaning for comprehenders (see also Traxler, 2011, 2012, 2014). Instead, the communicative context in which the sentence is situated ultimately helps make a sentence's meaning available to comprehenders. In most natural conversations, the context of an utterance reinforces the interpretation generated from the sentence's underlying structure, which allows people to comprehend utterances rapidly and accurately in real time. However, when comprehenders are deprived of the context, or if the available contextual information interferes with the interpretation generated from the underlying structure, then the normal routines of comprehension are disrupted, resulting in representations that are merely "good enough."

Ferreira and Patson (2007) formalize this notion in what they call the Good Enough approach to human language comprehension, founded on the view that the language processing system, by design, builds representations of utterances that are simply "good enough" for the task at hand (e.g., comprehension in conversation). For example, when the processing of a sentence becomes costly, comprehenders may resort to constructing a shallow representation of the sentence that may sometimes conflict in interpretation with the one generated from the faithfully parsed syntactic structure. In this way, the Good Enough approach attempts to unify the phenomena of comprehension errors beyond just those involving sentences with local syntactic ambiguity, such as thematic role reversals in passives and Moses illusions (Ferreira, 2003).

However, there are strong reasons to maintain the belief that the parser is capable of building complex structures under pressure, and that the syntactic structure still plays a significant role in determining the possible interpretations that are made available to comprehenders. Decades of psycholinguistics research on the role of grammatical constraints in parsing have found that comprehenders are sensitive to the structural relations between abstract linguistic units in a sentence that cannot be represented at a shallow level of processing. For example, readers have been shown to respect island constraints when positing gaps (Stowe, 1986; Traxler & Pickering, 1996; Phillips, 2006) and Binding Principles when forming pronoun-antecedent dependencies (Sturt, 2003; Kazanina et al., 2007; Kush & Dillon, 2021) in real time, neither of which are easily reducible to lexical identity, word order, pragmatic inference, and the like. In fact, even in the case of passives, which are often claimed to be especially susceptible to interference from extra-syntactic strategies like the agent-first heuristic (e.g., Ferreira, 2003), recent evidence suggests that the comprehension difficulty is driven instead by task-specific demands giving rise to “active-like” representations in offline processing, and the robustness of these interpretations to memory decay (Grillo et al., 2018; Paolazzi et al., 2019). Given this evidence, the claim from early formulations of the Good Enough approach – that the syntactic structure can become a fragile source of meaning for comprehenders when the sentence is too complex – warrants greater scrutiny.

This issue was investigated in Slattery et al.’s (2013) study of the syntactic representations that underlie “good enough” interpretations. Slattery et al. (2013) lay out two competing hypotheses about the nature of post-repair syntactic representations when misinterpretations are observed. The first hypothesis is that the parser fails to build the correct structure in reanalysis, resulting in a degenerate structure that tolerates misinterpretations. The second hypothesis is that the correct structure is fully represented, but the locally coherent analysis lingers in some form and interferes with comprehension. Early proposals of the Good Enough approach (e.g., Christianson et al., 2001; Ferreira et al., 2001) are most consistent with the first possibility, where an ill-formed and/or underspecified syntactic representation can drive comprehenders to become more reliant on external heuristics that operate quickly at the surface level of the sentence. However, if the correct syntactic structure is fully recovered after undergoing reanalysis, that presents a potential challenge to this account: it would need to make an even stronger claim about the ability of surface-level heuristics to dominate structural information as the primary source of sentence meaning (Fujita, 2021).

Slattery and colleagues tested both possibilities in two separate eye-tracking-while-reading experiments. In the first experiment, they exploited the gender mismatch effect (Sturt, 2003) as a diagnostic for whether the parser successfully builds the correct syntactic structure, using a pair of sentences such as in (6).

- (6) After the bank manager telephoned David's **father/mother** grew worried and gave himself approximately five days to reply.

In the absence of the post-adjunct comma, the classic garden path effect is predicted at the matrix verb *grew* for both sentences. But critically, if the parser is able to recover the correct structure after undergoing reanalysis, an additional difficulty at the pronoun *himself* is predicted when the matrix subject does not agree in gender (i.e., when the matrix subject is *David's mother*) because the parser's search for the antecedent of the reflexive should be constrained by the c-command constraint on reflexive binding (Chomsky, 1981). In other words, if the post-repair syntactic representation minimally specifies the c-command relationship between the matrix subject and the reflexive, and if the matrix subject is *David's mother*, then the parser will attempt to interpret *himself* as *David's mother* and experience difficulty doing so. On the other hand, if reanalysis is "good enough" and simply yields a coarse-grained representation of the sentence such that the antecedent search is not syntactically constrained, then the mere presence of the noun *David* preceding the reflexive pronoun should be a sufficient cue to erroneously license the interpretation of *David* as the antecedent of *himself*. Interestingly, a slowdown at the pronoun was observed only when the pronoun did not agree in gender with the matrix subject, suggesting that the parser does indeed build rich and detailed representations of the syntactic structure even when comprehension appears to be degraded.

At this point, if we still maintain the assumption that interpretations are constrained by the syntactic structure, then the parser must have built a structure that can license a misinterpretation at some point during processing, in addition to ultimately building the globally correct structure after reanalysis. This assumption bears out in Slattery et al.'s (2013) second experiment, which investigated whether the post-repair interpretation is consistent with that associated with the locally coherent parse by manipulating the material past disambiguation, such as in (7).

- (7) While Frank dried off the truck that was dark green was peed on by a stray dog. Frank quickly finished drying himself off then yelled out the window at the dog.

The adjunct verb *dried off* in the first sentence of (7) is temporarily ambiguous in transitivity until the matrix verb *was (peed on)* signals that the NP headed by *truck* is not the embedded object – i.e., that the transitive analysis of *dried off* is no longer viable. If that incorrect analysis is nevertheless available after disambiguation, comprehenders should experience difficulty processing *himself* in the following sentence, since the interpretation that Frank has finished drying himself off presupposes that he had been drying himself off before – an interpretation that is only available if the matrix verb *dried off* in the preceding sentence had incorrectly been interpreted as a reflexive at some point. Indeed, readers slow down at the pronoun well after the garden path effect in the first sentence had subsided, suggesting that some aspect of the initial misanalysis is persisting past the point of disambiguation.

Thus, Slattery et al. (2013) provide strong evidence that it is not the failure to construct the correct structure but rather the failure to discard the incorrect analysis that leads to lingering misinterpretations. This is consistent with previous findings for the semantic persistence effect as described in Sturt (2007), which studied reading behavior in garden path sentences when they followed a context sentence. Sturt reports that given a context sentence like (8), late reading measures indicative of semantic integration were longer at the critical disambiguating region (*was actually*) when the content of the second sentence was semantically inconsistent (8b) compared to when it was consistent (8a).

- (8) The Antarctic expedition had been going on for months.
- a. The explorers found the South Pole was actually right at their feet.
  - b. The explorers found the South Pole was actually impossible to reach.

The increased difficulty in processing (8b) relative to (8a) indicates that the initial semantic interpretation from the preceding context sentence must have been available at the point of reanalysis to interfere with semantic integration. Critically, this finding points to the possibility that this same mechanism of semantic persistence across two sentences may also characterize the lingering effect observed for garden path sentences before and after reanalysis as well.

To address this question of whether it is the *semantic interpretation* or the *syntactic structure* that lingers from the incorrect local parse to interfere with semantic/syntactic reanalysis, previous research has turned to the contribution of lexical factors. For garden path sentences with early/late-closure ambiguity like *While Anna dressed the baby played in the crib*, two qualities of the erroneous local analysis have been reported to contribute to the failure of reanalysis: the *semantic fit* of the adjunct NP as the direct object of the adjunct verb (Trueswell & Tanenhaus, 1994; Pickering & Traxler, 1998; Tabor et al., 2004; Hare et al., 2009) and the *transitivity bias* of the verb (MacDonald, 1994; MacDonald et al., 1994; Garnsey et al., 1997; Gibson & Pearlmuter, 1998, 2000). For example, if *the baby* is a plausible candidate as the object of *dressed* and if *dressed* is frequently used in transitive structures, comprehenders would have greater difficulty accurately answering a question that asks whether Anna dressed the baby.

However, this line of evidence has not been successful in disentangling the role of semantic vs. syntactic reanalysis in the lingering effect. For example, proponents of serial models with incremental semantic processing (e.g., Sturt, 2007) claim that the incorrect local analysis is first pursued before the globally correct analysis is made available to the parser, and so it is merely the semantic interpretation from that initial misanalysis which lingers. Under this account, the semantic processor may make stronger commitments to the initial misanalysis if the structure is highly probable and the interpretation is highly coherent, strengthening its resistance to revision during semantic reanalysis. On the other hand, proponents of parallel models that support the coexistence of multiple local structures (Tabor & Hutchins, 2004; Tabor et al., 2004; Ferreira et al.,



2004; Lau & Ferreira, 2005) claim that a locally coherent parse that was temporarily entertained could fail to decay if it receives sufficient activation from plausibility and frequency. Under this account, the lingering of the incorrect structure alongside the correct structure in the syntactic representation is what may drive the misinterpretations, despite the parser having successfully built the correct structure. Given this difficulty of disentangling the role of semantic versus syntactic reanalysis from the contribution of lexical information, we propose in the next section a method of isolating the role of semantic reanalysis in the lingering effect by manipulating post-lexical prosodic information, drawing on the well-known relations between pitch accent, focus, and semantic processing.

## 2.2 Pitch accent and meaning

In the Autosegmental Metrical (AM) model of English intonational phonology, the pitch contour of an utterance is understood to be derived from interpolation between discrete tonal targets which are anchored to syllables that are phrase-level peaks of prominence and phrase-final edge tones (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Gussenhoven, 2004). Phrasal prominence is determined from a hierarchically organized metrical structure, which defines a strong-weak patterning over words at the phrase level (Ladd, 2008). In English, the default prominence pattern has an obligatory prominence on the rightmost content word in the prosodic phrase. For a simple sentence (with no embedded clause), produced in a neutral speaking style, the verb and its (non-clausal) complements constitute one prosodic phrase, and a non-pronominal subject NP may optionally constitute a separate prosodic phrase, or may be integrated into the phrase containing the verb (Büring, 2016). An obligatory prominence at the right edge of a prosodic phrase marks the location of the main phrasal stress and is termed the *nuclear* prominence (Chomsky & Halle, 1968). If a phrase is long enough, there may be additional optional *prenuclear* prominences assigned to words in metrically strong positions earlier in the phrase (Calhoun, 2006).

A word with phrasal prominence licenses a tonally specified *pitch accent* (Truckenbrodt, 1995), which is obligatorily assigned to the word with nuclear prominence, and is optionally assigned as a “rhythmic” or “ornamental” feature of words with prenuclear prominence (Calhoun, 2006; Ladd, 2008; Büring, 2016). In the AM model, pitch accents may consist of a single tone (L\* or H\*), or a sequence of two tones (e.g., L+H\*), associated with the syllable with word-level primary stress in the prominent word. Additional tones are assigned at the right edge of a prosodic phrase, with a single tone (L- or H-) at the edge of a low-level prosodic phrase which is followed by a second tone (L% or H%) at the end of a high-level prosodic phrase. The following example illustrates the default hierarchical metrical structure, accent assignment, and phrase-edge tones for the sentence *Chicago is a city in Illinois*.





- (11) A: What did John eat for dinner?  
B: He ate CAKE.

Experimental works on prosodic focus over the last several decades provide supporting evidence of a systematic relationship between pitch accents and focus, both in production and comprehension (see reviews in Cutler et al., 1997; Wagner & Watson, 2010; Wagner, 2020; but see also Chodroff & Cole, 2019 and Roettger et al., 2019 for evidence of a weaker relationship). For example, studies using eye-tracking in the visual world paradigm have found that listeners can rapidly access information-structural meanings from pitch accents in complex comprehension tasks (Ito & Speer, 2008; Kurumada et al., 2014). Furthermore, interpretations of focus-related semantic meaning have been found to trigger deeper semantic processing that has consequences for encoding and retrieval in memory. In a series of recognition tasks, Fraundorf et al. (2010) found that listeners were better able to recall a word from a prior discourse when the word was accentually prominent. Critically, listeners performed better on the recognition task when the choice was between contrasting referents, meaning that contrast can be encoded in pitch accents in a way that facilitates future identification and discrimination.

These findings indicate that listeners rapidly and automatically compute the rich semantic and pragmatic meanings of pitch accents in sentence processing. However, as already noted, not all pitch accents mark semantic focus on the accented word. Rather, prosodically licensed semantic focus scopes over syntactic constituents in a derivational process known as focus projection (Gussenhoven, 1992, 1999; Selkirk, 1995). Revisiting the example in (11), because the most prominent pitch accent in Speaker B's answer (by virtue of being the only one present) occurs on *cake*, it specifically marks that object noun phrase as the answer focus. The focused constituent (a.k.a. the domain of focus) is marked with  $[\ ]_{\text{FOC}}$  brackets in (12), borrowing from Selkirk (1995). It is due to this ability for focus to project from the accented word and percolate up the syntactic structure that the same response by Speaker B in (12) is also an appropriate answer to a VP-focus question like (13).

- (12) A: What did John eat for dinner?  
B: He ate  $[\ ]_{\text{FOC}}$  CAKE  $[\ ]_{\text{FOC}}$ .

- (13) A: What did John do?

Theories of the intonational grammar of American English generally agree that, to mark focus, the last pitch accent in the focus domain must bear the nuclear pitch accent and all subsequent material in the sentence must be deaccented (e.g., Jackendoff, 1972; Schwarzschild, 1999; Ladd, 2008). This focus-to-accent relationship becomes clearer in sentences where focus and default accenting locate the nuclear pitch accent on different parts of the utterance. For instance, the default accenting of (14) assigns prominence to *CLASS*, the rightmost word in the intonational

phrase eligible to receive a pitch accent. However, having a contrastive focus over *Suzy* in (14) requires that the nuclear pitch accent occurs within the domain of focus, which is realized in (15) with an accent on *Suzy*. In turn, the accent on *class* is removed, as per the rule of post-focal deaccenting.

(14) Suzy showed up to CLASS.

(15) No, not John. [SUZY]<sub>FOC</sub> showed up to class.

That the domain of focus must minimally contain the nuclear pitch accent is also reflected in the observation that focus is not realized by simply placing a pitch accent just anywhere in the domain of focus. For example, given Speaker A's question in (16), the verb phrase *visited a friend in Chicago* in Speaker B's answer constitutes the domain of focus. This may be realized prosodically with an H\* on *Chicago*, as in (16a), but it would be inappropriate to have the nuclear pitch accent occur on any other word in the focus domain, such as on *visited* (16b). Since *visited* is phrase-medial and a predicate preceding an argument, assigning it nuclear (i.e., rightmost) prominence makes available only the narrow verb-focus reading, as shown in (17).

(16) A: What did you do last weekend?

a. B: I [<sub>VP</sub> visited a friend in CHICAGO]<sub>FOC</sub>.

b. B: \*I [<sub>VP</sub> VISITED a friend in Chicago]<sub>FOC</sub>.

(17) I [<sub>V</sub> visited]<sub>FOC</sub> a friend in Chicago.

Given these constraints on the interpretation of prosodic focus and the observations on the effect of pitch accent on semantic processing in real time, we can make predictions about the interpretation of prominence in more complex sentences. Consider again the sentence *While Anna dressed the baby played in the crib*. Over the course of processing this sentence, the parser is expected to have built both the incorrect and the correct structure – either simultaneously (parallel) or in succession (serial), depending on the theory – due to the preference for late closure. Critically, the incorrect local analysis that was temporarily entertained may persist depending on the strength of local coherence as noted earlier. Here, we posit that the likelihood of persistence may also be affected by prosodic factors as well.

For example, the sentence can be produced with a sharply rising pitch accent on the adjunct verb *dressed*, as in *While Anna DRESSED the baby played in the crib*. When the parser is garden-pathed and initially interprets *the baby* as the object of *dressed*, the accent on the adjunct verb occupies a non-canonical, phrase-medial position in that structure. Assuming that the parser considers a focus interpretation when there is structural evidence to do so in the absence of relevant context, the phrase-medial pitch accent is strong evidence for the presence of narrow

focus on the transitive interpretation of the verb.<sup>3</sup> This assumption generates an interesting prediction for a serial parser with incremental semantic processing: the prominence on the adjunct clause verb should yield an interpretation of focus on the incorrect late-closure parse, triggering deeper semantic processing for the erroneous analysis and enriching it with focus-related semantic meaning before reanalysis is triggered.

A likely arrangement of ordinary semantic representations (Carlson, 1984) for the early/late-closure parses in each accent condition is illustrated in (18), with the region of temporary local syntactic ambiguity in bold. If the pitch accent is interpreted as narrow focus (in 18a and 18c-d), the focus semantic meaning is also provided, where focus contributes to the semantic value with a set of alternatives. Following Rooth's (1992) Focus Interpretation Principle, focus is understood to introduce a free variable at the level of the domain of focus, bound through the existential closure as an argument.<sup>4</sup>

- (18) a. While Anna [<sub>VP</sub> [<sub>V</sub> **DRESSED**]<sub>FOC</sub> **the baby**], ... (verb accent, late closure)  
*Ordinary semantic:*  $\exists e[\text{dress}(e) \ \& \ \text{Agent}(e, \text{Anna}) \ \& \ \text{Theme}(e, \text{the baby})]$   
*Focus semantic:*  $\exists e[P(e) \ \& \ \text{Agent}(e, \text{Anna}) \ \& \ \text{Theme}(e, \text{the baby})]$
- b. While Anna [<sub>VP</sub> **DRESSED**], **the baby** ... (verb accent, early closure)  
*Ordinary semantic:*  $\exists e[\text{dress}(e) \ \& \ \text{Agent}(e, \text{Anna})]$   
*Focus semantic:*  $\emptyset$
- c. While [<sub>NP</sub> ANNA]<sub>FOC</sub> **dressed the baby**, ... (subject accent, late closure)  
*Ordinary semantic:*  $\exists e[\text{dress}(e) \ \& \ \text{Agent}(e, \text{Anna}) \ \& \ \text{Theme}(e, \text{the baby})]$   
*Focus semantic:*  $\exists x[A(x)]$  (denotes the alternative set of {Anna'})
- d. While [<sub>NP</sub> ANNA]<sub>FOC</sub> **dressed, the baby** ... (subject accent, early closure)  
*Ordinary semantic:*  $\exists e[\text{dress}(e) \ \& \ \text{Agent}(e, \text{Anna})]$   
*Focus semantic:*  $\exists x[A(x)]$  (denotes the alternative set of {Anna'})

Here, we see an asymmetry in the interpretation of focus between the local parses within the verb accent condition (18a-b), compared to within the subject accent condition (18c-d). Whereas the late-closure parse in the verb accent condition is enriched with focus semantic meaning which must later be revised (18a), the focus semantic meaning from the late-closure parse in the subject accent condition (18c) need not be revised in the course of garden path reanalysis, because a

---

<sup>3</sup> As pointed out by a reviewer, it is possible for the pitch accent on the verb to be interpreted as broad focus, projecting to a larger unit than the verb such as the VP or the sentence. However, given its marked position (a predicate preceding its argument in the same prosodic phrase), a narrow-focus interpretation is most consistent with this verb accent contour in the late-closure parse. Many formal models of focus projection explicitly reject (Selkirk, 1995; Gussenhoven, 1999) or place high-ranking constraints on (Truckenbrodt, 1995; Schwarzschild, 1999; Büring, 2007) the projection of focus from V to VP (and to any larger unit) when the internal argument is not accented and the sentence is all-new.

<sup>4</sup> In the focus semantic meaning of (18), *P* and *A* are free variables representing alternatives to *dressed* and *Anna* (i.e., the alternative sets {dressed'} and {Anna'}).

focused subject (*Anna*) is equally compatible with the early- and late-closure interpretations. In other words, the focus may in fact not be a target of reanalysis entirely (since it lies outside the region of ambiguity). Thus, we predict no additional semantic processing for the erroneous analysis in the subject accent condition, assuming that the interpretation of the pitch accent on the subject is unaffected by the process of ambiguity resolution which follows it. In this sense, the subject accent contour serves as the baseline condition that allows us to isolate the effect of deeper semantic processing on the initial misanalysis from the lower-level lexical effects on local coherence. Additionally, although (18c-d) construes the accent on the subject as contrastive, it is also possible for the subject accent contour to be interpreted as a rhythmic or “ornamental” accent, i.e., unrelated to focus (Büring, 2007). But whatever the interpretation of the subject accent may be, our hypothesis makes the same prediction: there should be no direct consequences for semantic reanalysis since the accented subject lies outside the region of ambiguity (in other words, there is no need to revise the initial interpretation of the subject accent). In that sense, the critical manipulation that separates the verb accent condition from the baseline subject accent condition is that the verb accent facilitates an interpretation of focus for the erroneous analysis *which must later be revised*, in order to be consistent with the syntactic reanalysis of the structure of the adjunct VP.

In sum, our hypothesis is as follows: if the lingering effect is driven by the failure of *semantic reanalysis* (i.e., the failure to discard the erroneous semantic interpretation), then enriching the erroneous semantic interpretation with focus-related semantic meaning should strengthen this effect. The crucial comparison here is comprehension accuracy between the condition where the erroneous garden path interpretation is marked for contrastive focus (verb accent condition) and the condition where it is not (subject accent condition).<sup>5</sup> If the pitch accent on the verb facilitates an interpretation of narrow semantic focus in the incorrect local parse, then the erroneous semantic interpretation will become more likely to linger than in the baseline condition with the pitch accent on the subject. We predict this lingering misinterpretation to be reflected in lower accuracy rates on comprehension questions that ask whether the second NP is the embedded object (e.g., *Was the baby dressed?*), an interpretation that is only licensed by the incorrect late-closure parse. Critically, if this effect of pitch accent on the depth of semantic processing drives the persistence effect independently of the extent of locally determined coherence, then the location of the pitch accent should emerge as a significant predictor of accuracy on comprehension questions after controlling for item-level lexical effects from semantic fit and transitivity bias.

---

<sup>5</sup> An interesting nuance to this hypothesis which is beyond the scope of this paper is whether the depth of semantic processing in the *correct* analysis (i.e., in the early-closure parse) also modulates the strength of semantic persistence. We take this issue to be separate here given our hypothesis that the semantic persistence effect is driven by the failure to discard the incorrect analysis, which does not directly speak to the role of the true analysis consistent with the syntax. Enriching the correct analysis with focus semantic meaning may facilitate the discarding of the incorrect analysis, or it may not have much of an effect since the lingering effect has been observed in reading studies even when the correct analysis was congruent. We leave this topic for future research.

### 3. Experiment

Two off-line norming experiments and a speeded auditory comprehension experiment were hosted on PCIBex (Zehr & Schwarz, 2018) and administered through Prolific, a platform for online subject recruitment. Thirty adult English monolingual speakers between 18 and 65 years of age residing in the United States were recruited for each of the two norming experiments, and sixty-one participants of the same demographic background (with the additional qualification of having no known hearing issues) were recruited for the speeded auditory comprehension experiment.

#### 3.1 Design and norming

The critical items for the auditory comprehension experiment consisted of twenty-four garden path sentences with closure ambiguity, such as (19). The adjunct clause consisted of a complementizer (*since*, *while*, or *when*) followed by a two-syllable name with a strong-weak stress pattern and an ambiguously transitive verb in the past tense. The matrix clause consisted of a subject noun phrase that was also a possible direct object candidate for the adjunct clause verb, followed by a matrix verb phrase.

(19) While Anna dressed the baby stopped crying.

An initial set of twenty-six sentences were normed for biases in *semantic fit* (how often the interpretation of the second NP as the embedded object is preferred) and *transitivity bias* (how often the adjunct verb is followed by a direct object) in two independent off-line experiments. The norming experiments are described in more detail as follows:

Semantic fit was normed in a three-alternative forced-choice task, in which participants were asked to choose the more plausible event between the early-closure interpretation such as (20), the late-closure interpretation such as (21), or an option for both being equally plausible. Participants in the semantic norming task did not rate locally ambiguous sentences such as (19).

(20) While Anna dressed, the baby stopped crying.

(21) While Anna dressed the baby, he stopped crying.

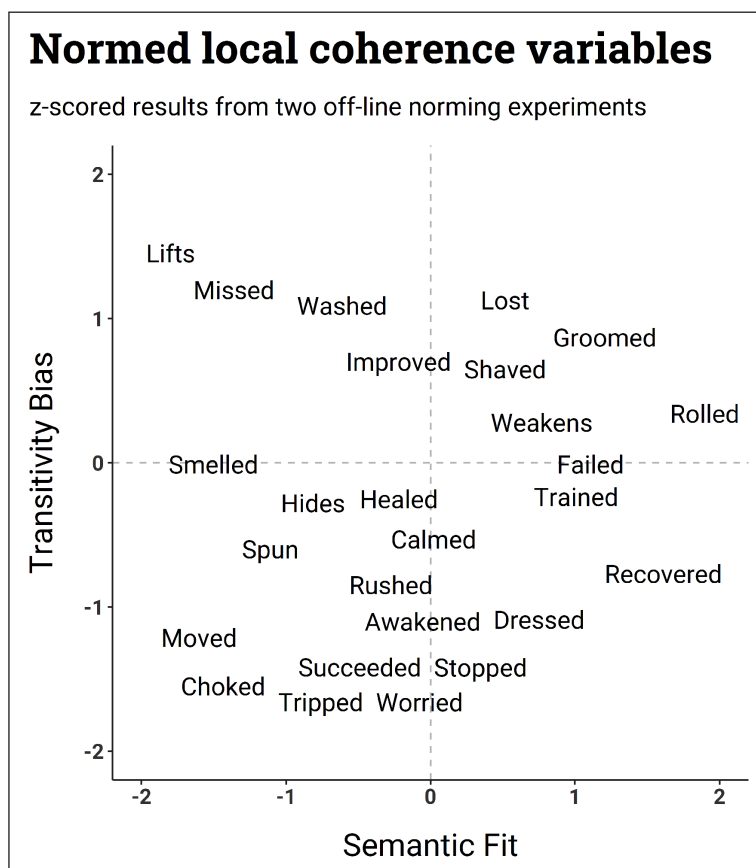
For each item, the proportion of responses preferring the early-closure interpretation against those preferring the late-closure interpretation was calculated (*median* = 0.53, *IQR* = 0.38). Semantic fit scores of each item were then centered at 0.5 (representing equal bias) and downweighted by the number of “both” responses.<sup>6</sup>

---

<sup>6</sup> Normalized semantic fit score = early-closure/(early-closure + late-closure) \* (1 - “both”/total).

Transitivity bias was normed in a sentence completion task, in which participants saw fragments of each sentence up to the adjunct clause verb, such as *While Anna dressed the baby* \_\_, and were asked to type in a continuation that completes the sentence. Responses that failed to complete the sentence were excluded from analysis and the rest were coded for the presence of a direct object following the adjunct clause verb. The proportion of responses involving direct objects was calculated (*median* = 0.38, *IQR* = 0.60) and again centered at 0.5.<sup>7</sup>

The centered measures of semantic fit and transitivity bias were then z-scored to be entered into the statistical model as independent item-based predictors.<sup>8</sup> **Figure 1** shows the distribution of semantic fit and transitivity bias scores for all items used in the experiment. The variability among items along both dimensions allows the model to statistically control for their effects on comprehension accuracy and isolate the effect of pitch accent placement.



**Figure 1:** Norming scores for transitivity bias and semantic fit of the sentences used as stimuli for the speeded auditory comprehension experiment.

<sup>7</sup> Normalized transitivity bias score = direct object/(direct object + no direct object)

<sup>8</sup> The measures were z-scored with respect to 0.5 as the center instead of the mean, as 0.5 represents equal bias.



### 3.2 Audio stimuli

The stimuli for the auditory perception experiment were recorded using a final set of twenty-four normed sentences.<sup>9</sup> A trained male speaker recorded each sentence with two accent patterns – one with prominence on the adjunct subject, such as (22), and another with prominence on the adjunct verb, such as (23).

(22) While ANNA dressed, the baby stopped crying. (subject accent, early closure)

(23) While Anna DRESSED, the baby stopped crying. (verb accent, early closure)

In the construction of the auditory stimuli, we considered the possibility that an interpretation of a prosodic phrase boundary after the adjunct verb may immediately resolve the local attachment ambiguity, much like the role of a comma in reading. The effect of an intervening intonational phrase boundary on blocking attachment is well-documented and may be signaled via a combination of a pause, pitch reset, and pre-boundary lengthening (Kjelgaard & Speer, 1999; Schafer et al., 2000). In order to make these cues ambiguous, so as to induce the garden path effect, a set of late-closure variants of the stimuli, such as (24) and (25), were also recorded by the same speaker. From the late-closure recordings, the durations of the words over the region of ambiguity that are potential targets of pre-boundary lengthening – the adjunct verb (*dressed*) and the post-verbal noun (*the baby*) – were measured and referenced in the resynthesis of the early-closure stimuli used in the experiment.

(24) While ANNA dressed the baby, he stopped crying. (subject accent, late closure)

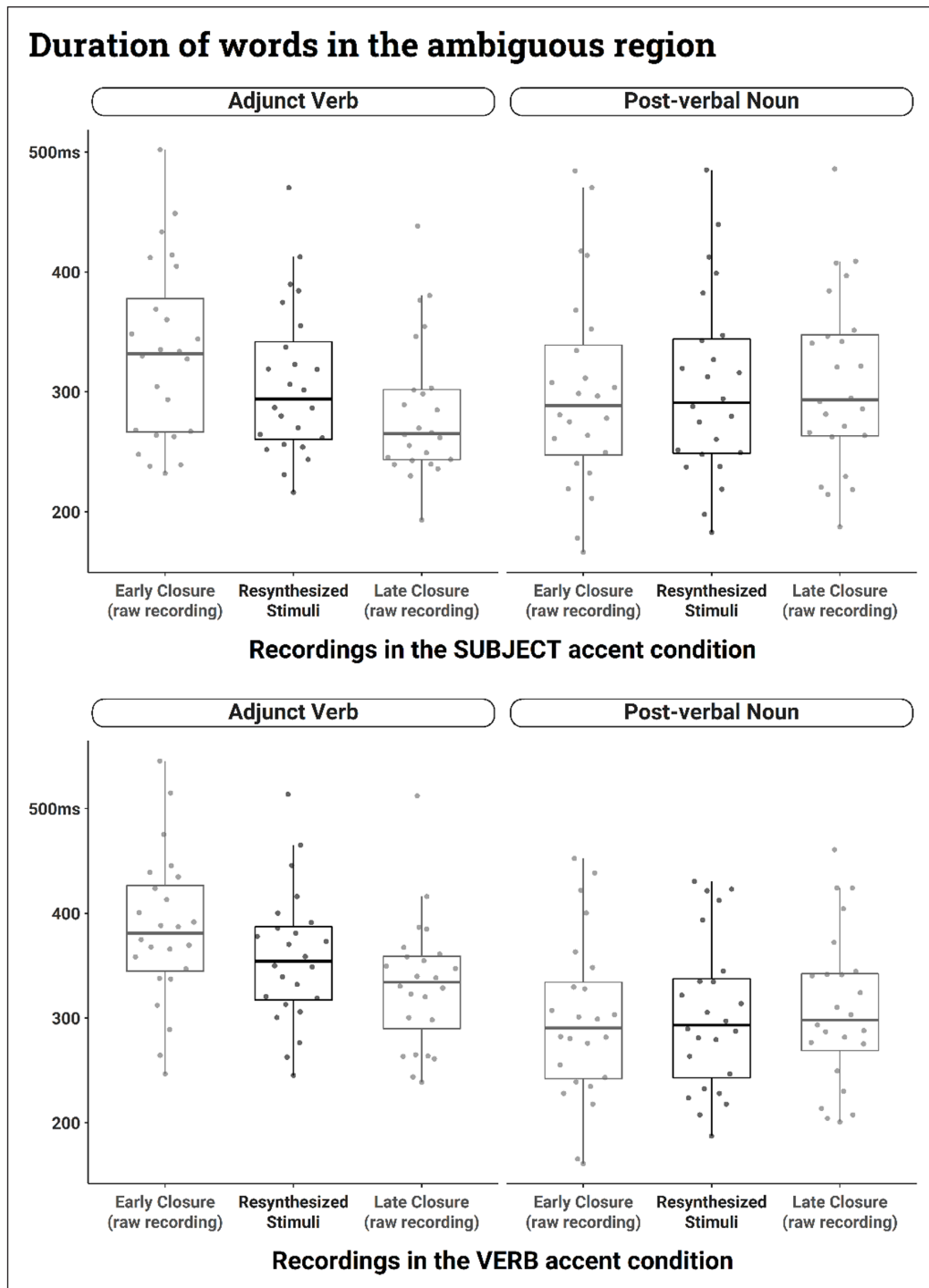
(25) While Anna DRESSED the baby, he stopped crying. (verb accent, late closure)

In total, four recordings were created for each sentence. **Figure 2** shows the distribution of durations for the two words that are potential candidates for pre-boundary lengthening in the four source recordings as well as in the resynthesized stimuli.

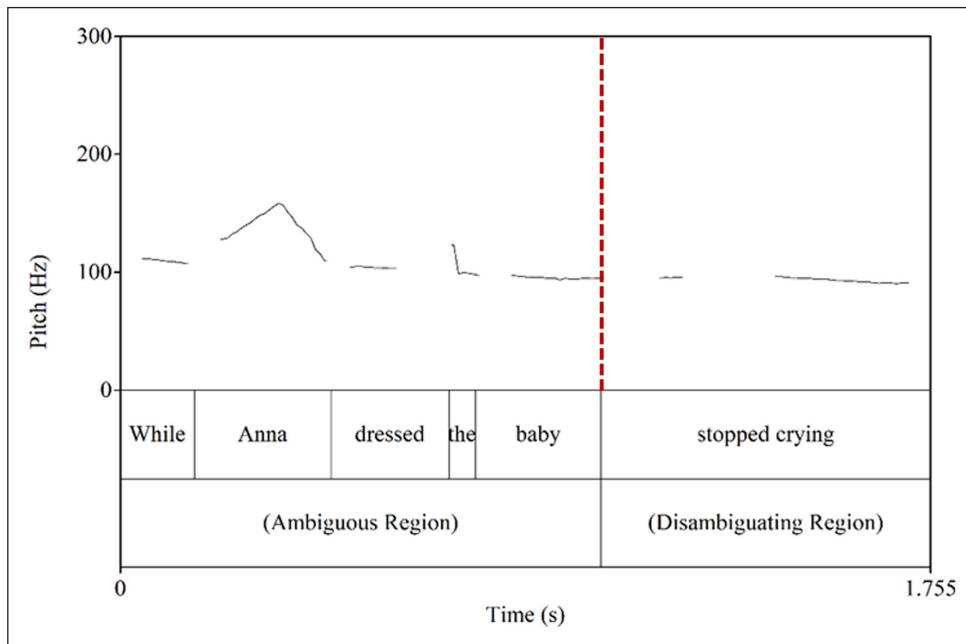
The duration of the adjunct clause verb and the post-verbal noun were resynthesized in Praat (Boersma & Weenik, 2020) to their averages between the early- and late-closure variants, using the early-closure recording as the source recording. After controlling for pre-boundary lengthening cues in this way, the end-of-sentence region (material after the post-verbal noun) from the subject-accent recordings was spliced into the verb-accent recordings to control for any sentence wrap-up effects between the two accent conditions. Lastly, pitch movement over the main clause was flattened to induce an unambiguous perception of accentual prominence in the adjunct clause. The pitch contours of the subject accent stimuli and the verb accent stimuli for a sentence used in the experiment are shown in **Figures 3** and **4** respectively, where the dotted line after the ambiguous region marks the location of the splice.

---

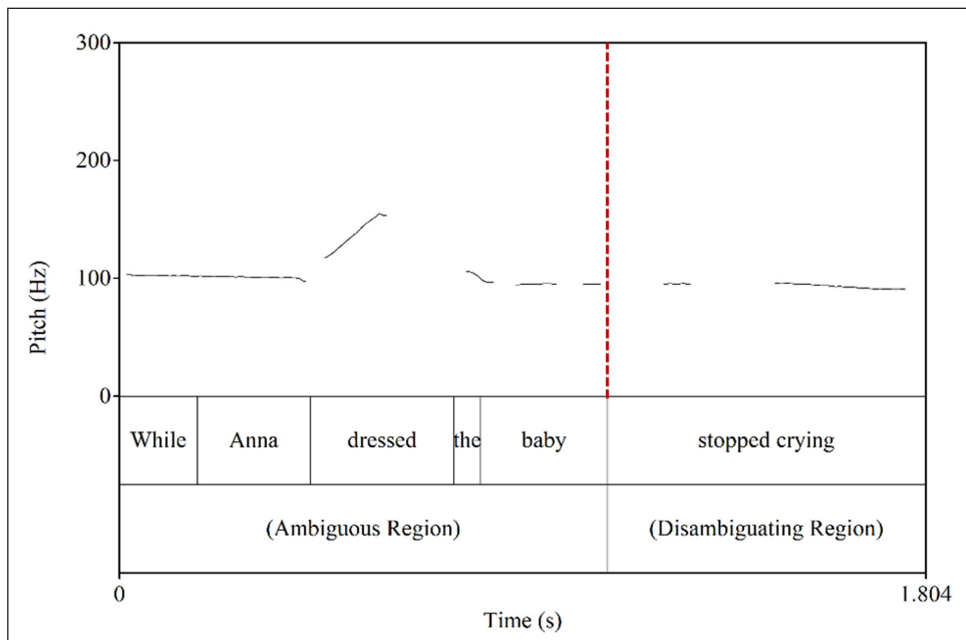
<sup>9</sup> Two items from the norming studies were removed to maximize spread of local coherence effects among the items for modeling purposes.



**Figure 2:** Distribution of durations (in milliseconds) for words in the ambiguous region that are candidates of pre-boundary lengthening: the adjunct verb (left) and the post-verbal noun (right). The duration distributions are plotted for recordings in the subject accent condition (top) and the verb accent condition (bottom). The source recordings are plotted in grey and the resynthesized stimuli which average between the source recordings are plotted in black and placed in the middle.



**Figure 3:** Pitch contour of the subject-accent stimuli for *While ANNA dressed the baby stopped crying*.



**Figure 4:** Pitch contour of the verb-accent stimuli for *While Anna DRESSED the baby stopped crying*.

Because prominence on a word has consequences for the prosodic structure of the entire phrase, acoustic measures between the conditions were not further controlled in order to preserve naturalness. In other words, the overall shape of the pitch contour over the ambiguous region was unaltered, aside from controlling for pre-boundary lengthening cues. **Figure 5** reports the measurements for pitch (max F0) on the adjunct verb and the post-verbal noun in both conditions from the resynthesized stimuli that were played to participants.

Condition	Word	Max F0		
		Mean	Std. Dev.	Distribution
Subject Accent	Adjunct Subject	155.2	9.5	
	Adjunct Verb	118.7	14.1	
Verb Accent	Adjunct Subject	109.0	8.1	
	Adjunct Verb	150.1	10.5	

**Figure 5:** The mean, standard deviation, and distribution of max F0 values on the adjunct subject and the adjunct verb for the resynthesized stimuli in the two accent conditions.

For each sentence, a corresponding comprehension question similar to that used in Christianson et al. (2001) was constructed. These questions asked whether the post-verbal noun was the direct object of the adjunct clause verb, such as in (26). Because the post-verbal noun in the critical sentences was always the subject of the matrix clause, the correct answer to all the comprehension questions for the critical trials was *NO*.

(26) Was the baby dressed? YES/NO

In addition to the critical items, forty-eight distractor sentences were recorded by the same speaker to balance the transitivity bias of the adjunct verb (ambiguous, obligatorily transitive, obligatorily intransitive), the location of prosodic prominence (first noun, first verb, second noun), and the sentence's syntactic structure (early- vs. late-closure of the adjunct clause VP) across all trials presented to participants in the experiment. The comprehension questions were also overall balanced in the correct answer choice and varied in the syntactic position of the entity being asked about. Lastly, two catch trials which simply instructed participants to answer *YES* or *NO* to the following question were included to filter out inattentive listeners.

As discussed previously, our hypothesis predicts a significant negative effect of verb pitch accent on comprehension accuracy. If a pitch accent on the verb leads to an interpretation of focus that triggers deeper semantic processing for the incorrect analysis, listeners should be more likely to answer the comprehension question in a way that is consistent with that erroneous

interpretation. For example, after listening to the sentence *While Anna DRESSED the baby stopped crying*, participants are predicted to incorrectly answer *YES* to the question *Did Anna dress the baby?* more frequently than in the corresponding utterance with an accent on the subject *ANNA*. Importantly, this effect of pitch accent location on accuracy should reach significance after controlling for transitivity bias and semantic fit. Given previous findings for these two factors contributing to the persistence of the merely local interpretation, we also expect a transitively biased adjunct verb followed by a plausible direct object candidate to conspire to strengthen the initial misanalysis, making misinterpretation more likely.

### 3.3 Procedure

Each experiment session began with two familiarization trials in which participants listened to two sentences with the same structure as the critical sentences, presented auditorily with no accompanying text.<sup>10</sup> Following the familiarization trials, the session continued with 74 trials, including 24 critical trials and 48 distractor trials, presented in pseudorandomized order that evenly spread out the critical trials. Each trial began with the audio presentation of a recorded sentence, with a blank screen, after which participants were given 5000ms to answer a comprehension question about the sentence that they just heard. The question was presented in written form and participants responded by clicking “YES” or “NO” on the screen. If an answer was not given within the time window, the next trial automatically proceeded. There were also two randomly sequenced catch trials which simply explicitly instructed participants to answer *YES* or *NO* to the following comprehension question. The experiment was run in a within-participant design where participants heard the critical sentences in twelve subject-accent and twelve verb-accent conditions. The experiment lasted around 10 to 15 minutes. The answers and the response times to the comprehension questions were recorded for analysis.

### 3.4 Results

Prior to analysis, critical trials where participants timed-out or answered the comprehension question in under 500ms were excluded (<2%).<sup>11</sup> Additionally, out of 61 participants, one was excluded due to low accuracy (<75%) on the distractor trials (mean accuracy on the distractor trials for the included participants was 95%). Data from the remaining sixty participants were included in the analysis.

The grand mean of within-speaker accuracies was above chance at 50% but well below ceiling in both the verb accent (78%) and subject accent (83%) conditions, replicating the semantic

---

<sup>10</sup> The exposure phase was also used to screen out participants with audio issues; every participant passed.

<sup>11</sup> The threshold was based on the approximate time it takes to read and understand the comprehension questions.

persistence effect. To test whether the difference between the two conditions was significant after controlling for the lexical-level predictors, a logistic mixed-effects regression model was fitted to the accuracy of the responses using the *lme4* package in R (Bates et al., 2015), with pitch accent condition, semantic fit, and transitivity bias as fixed effects without interaction terms, and subject and item as random effects.<sup>12</sup> Convergence issues due to overparameterization of the random effects structure were addressed through the iterative model reduction process suggested in Bates et al. (2018), whereby the correlation between random effects was dropped first and individual random effects were removed in the order of lowest variance until the model converged, in order to keep the random effects structure as informative as possible. The final model included a random slope for Pitch by subject, and random intercepts by item and by subject.<sup>13</sup> To test the significance of each predictor, the full model was compared via a log likelihood-ratio test to a depleted model with the predictor of interest removed.<sup>14</sup> **Table 1** reports the log-odds estimates and standard errors from the final model and the chi-squared statistics and p-values from the log likelihood-ratio tests.

**Table 1:** The point estimate and standard error (in brackets), chi-squared value, and the p-value (bold for reaching significance,  $p < 0.05$ ) from the logistic mixed-effects model fitted to accuracy.

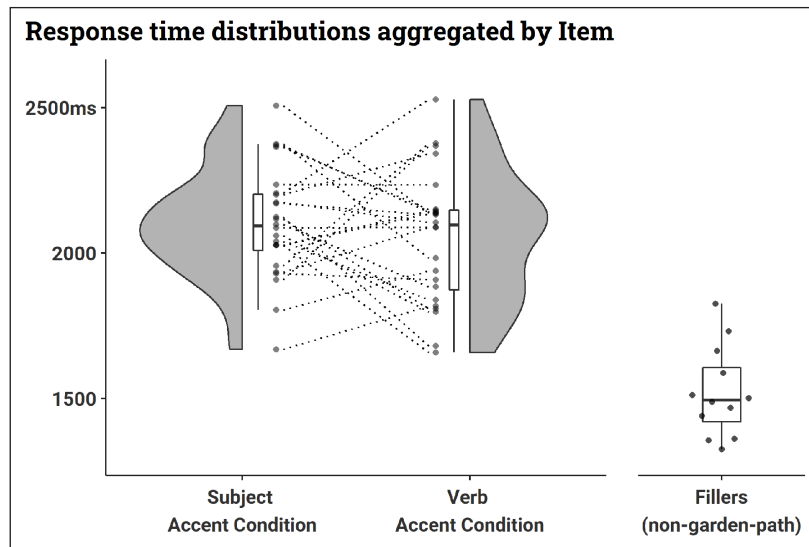
	<i>Estimate (SE)</i>	$\chi^2$	<i>p</i>
Pitch (Verb)	−0.19 (0.08)	5.41	<b>0.020</b>
Semantic Fit	−0.44 (0.15)	7.45	<b>0.006</b>
Transitivity	−0.19 (0.16)	1.39	0.238

A post-hoc analysis of response times ( $mean = 2080$ ,  $SD = 759$ ) was also conducted to determine whether listeners were garden-pathed in both pitch accent conditions, as intended by the resynthesis design. **Figure 6** shows the distribution of response times aggregated by item between the two conditions, where each line tracks an item’s mean response time across conditions. The response time distribution for a subset of the filler items that were non-garden-path sentences of similar length with an unambiguously intransitive adjunct verb (e.g., *When Tyler sneezed, the driver passed a tissue*) is also plotted to the right for comparison.

<sup>12</sup> Accuracy  $\sim$  PitchAccent + SemanticFit + TransitivityBias + (1 + PitchAccent + SemanticFit + TransitivityBias | Subject) + (1 + PitchAccent | Item).

<sup>13</sup> Accuracy  $\sim$  PitchAccent + SemanticFit + TransitivityBias + (1 + PitchAccent | Subject) + (1 | Item).

<sup>14</sup> For example, with Pitch Accent removed: Accuracy  $\sim$  SemanticFit + TransitivityBias + (1 + Pitch Accent | Subject) + (1 | Item).



**Figure 6:** Response time distributions of the critical trials by condition (raincloud plot to the left) and a subset of the filler trials (boxplot to the right). Each dot represents the mean response time of an item. The slopes of the lines connecting the dots represent the effect of pitch accent condition on response time.

In addition, a linear mixed-effects regression model was fitted to log-transformed response time with pitch accent placement, semantic fit, transitivity bias, and accuracy as fixed effects without interaction terms, and subject and item as random effects with varying intercepts.<sup>15</sup> **Table 2** reports the estimates, standard errors, and  $t$ -values from the model. The model estimates a large effect on response time of semantic fit and transitivity bias such that responses are slower for items with high semantic fit and high transitivity bias scores, whereas the location of the pitch accent has a small effect on response time.<sup>16</sup>

**Table 2:** The point estimate and standard error (in brackets), and the  $t$ -value from the linear mixed-effects model fitted to response time.

	<i>Estimates (SE)</i>	<i>t</i>
Pitch (Verb)	−0.009 (0.009)	−1.07
Semantic Fit	0.031 (0.014)	2.22
Transitivity	0.035 (0.015)	2.43
Accuracy	0.065 (0.021)	2.89

<sup>15</sup>  $\text{LogResponseTime} \sim \text{PitchAccent} + \text{SemanticFit} + \text{TransitivityBias} + \text{Accuracy} + (1 | \text{Subject}) + (1 | \text{Item})$ .

<sup>16</sup> Converted to probability scale, the liberal estimate of the effect of pitch accent from the 95% confidence interval  $[-0.026, 0.008]$  roughly corresponds to a 2.5% difference in response times between the two conditions. Assuming the average response time of 2080ms, this amounts to a difference of 54ms. For comparison, the difference in response times between the duration/length-matched unambiguous fillers and all critical items was roughly 500ms.



## 4. Discussion

The auditory comprehension experiment investigated the independent effect of the depth of semantic processing on the persistence of the erroneous, locally coherent interpretation in garden path sentences with early/late-closure ambiguity. The depth of semantic processing was manipulated through the prosodic marking of contrastive focus, under the assumption that a constituent with contrastive focus undergoes a deeper level of processing that includes identification of semantic alternatives to the focused constituent (i.e., the constituent that binds the free variable introduced by focus, not just the accented word itself; Rooth, 1992). The critical prosodic manipulation involved contrastive focus marked by a sharply rising pitch movement on the adjunct clause verb, which facilitates an interpretation of contrastive focus on the verb in the incorrect late-closure parse due to the marked, phrase-medial position of the accent in that structure. This contrastive focus in the incorrect parse was predicted to enrich the semantic value of the transitive analysis with a set of alternatives to the predicate, thereby triggering deeper semantic processing and strengthening its resistance to revision at reanalysis. In contrast, a pitch accent marking focus on the subject of the adjunct clause was predicted to facilitate an interpretation of narrow focus, contributing to the semantic value of the sentence with a set of alternatives to the subject. Because the pitch accent and focus on the subject is processed at an earlier position in the sentence that is not the target of reanalysis (i.e., lies outside the region of ambiguity), we predicted no such additional difficulty for semantic reanalysis in the subject accent condition.

The results show a significant negative effect of verb pitch accent on accuracy ( $\beta = -0.19$ ,  $SE = 0.08$ ,  $\chi^2 = 5.41$ ,  $p = 0.02$ ), providing evidence that the processing of focus-related semantic meaning that contributes to the incorrect analysis strengthens semantic persistence. Critically, this effect is significant after controlling for item-specific local coherence effects from semantic fit and the transitivity bias. Therefore, the data is consistent with our hypothesis that pitch accents have independent semantic consequences in the processing of syntactically ambiguous material. This complements previous findings on the rapid and incremental interpretation of focus-related semantic meaning signaled by prosodic prominence in referential ambiguity resolution tasks using the visual world paradigms (Ito & Speer, 2008; Watson et al., 2008; Kurumada et al., 2014). We find evidence for such incremental processing of alternatives at the level of the focused constituent when the contribution of the focus is restricted to just the erroneous garden path interpretation, strengthening semantic persistence.

One may wonder whether there is an alternative explanation for the effect of pitch accent on accuracy based on the pitch accent *forcing* the parser to commit to a particular structural analysis. For example, the effect of pitch accent location on accuracy could be that, for whatever reason, the verb accent condition leads the parser down the garden path but the subject accent condition does not. If the issue is simply about whether the parser was garden-pathed in this sense, we would expect a strong corresponding effect of pitch accent on response time as

well, given the well-attested causal relationship between being garden-pathed and increased processing time (Frazier & Rayner, 1982; Kjelgaard & Speer, 1999). Specifically, we would predict shorter response times in the subject accent condition that are similar to the response times for the non-garden-path filler trials if the parser pursued the correct analysis from start to finish without encountering any ambiguity. Instead, as shown in **Figure 6** and the post hoc model in **Table 2**, we find that both pitch accent conditions have similar response times and are longer than the non-garden-path distractor trials. This is consistent with the interpretation that the parser was garden-pathed regardless of the location of the pitch accent in the adjunct clause, as expected given that the resynthesized stimuli had ambiguous phonetic cues to a prosodic phrase boundary over the region of attachment ambiguity. Of course, since response time is an imprecise measure of the time-course of parsing, the location of the pitch accents may in fact have syntactic consequences that are drowned in the noisiness of the measure. But assuming that participants were indeed garden-pathed in all critical trials, our findings for accuracy likely reflect failures in the process of semantic reanalysis, as opposed to difficulties with syntactic structure building and reanalysis.

At this point, we have presented evidence for the following two claims. First, misinterpretations are driven by interference from the semantic interpretation associated with the incorrect structural analysis, the strength of which can be modulated by prosodically marked, focus-related semantic meaning. Second, independent of this effect of pitch accent placement on comprehension, the local syntactic ambiguity is still present at the post-verbal noun, and the parser must eventually build the structure that is faithful to the input. These conclusions from the data are broadly consistent with the account of semantic persistence as proposed in Sturt (2007) and Slattery et al. (2013).

With regards to the issue of whether it is the syntactic structure or just the semantic representation that persists, findings from the response time analysis are more consistent with the latter. This is because the former predicts that the processing of semantic focus in the incorrect local parse will increase the activation level of that parse as a whole, making the overall process of ambiguity resolution more costly. For example, under the assumption that “successful comprehension is about successfully integrating new information,” Ferreira & Lowder (2016, p. 239) claim that the comprehension system specifically targets new and focused information and gives it an advantage in processing. If the parser prioritizes structures that contain focused information in the allocation of attentional resources during ambiguity resolution in the parallel architecture, the incorrect parse should become a stronger competitor to the globally correct parse, resulting in not only lower accuracy but also longer response times. Instead, results are more consistent with an account where pitch accents are simply interpreted with respect to each parse that is being considered before and after syntactic reanalysis. For example, in the verb accent condition, the presence of the early/late-closure ambiguity and the preference for local

attachment leads the parser to first construct the late-closure analysis of the adjunct VP, within which the pitch accent is interpreted as narrow focus on the transitive interpretation of the verb. When the disambiguating material downstream triggers reanalysis and the parser builds the correct early-closure parse, the pitch accent on the verb may be *reinterpreted* with respect to this new structure. Importantly, this means that the success of syntactic reanalysis is independent of the success of semantic reanalysis, such that the failure of semantic reanalysis can affect the accuracy of comprehension without necessarily affecting the time course of building the correct structure.<sup>17</sup>

The semantic processor's failure to discard the initial semantic commitments to the incorrect parse despite the success of syntactic reanalysis in this way can be best captured in a serial parsing architecture that also allows some degree of incremental semantic processing. Sturt's (2007) proposal that syntactic reanalysis always succeeds (i.e., the parser never builds ungrammatical structures) but the semantic processor can independently fail to revise its initial commitments is one possible account of our findings. The sensitivity of the semantic processor to the focusing of an incorrect local analysis suggests that a pitch accent on a syntactically ambiguous word may be immediately interpreted with respect to the syntactic structure of the parse(s) being considered, at least in the absence of clear phonetic cues to a prosodic phrase boundary. Such a possibility for semantic reanalysis to fail and allow a misanalysis to persist undetected has been explored in the processing of other types of ambiguities as well (see reviews in Sanford & Sturt, 2002). For example, Baker and Wagner (1987) found that although the semantic anomaly in the so-called "Moses illusions" (e.g., *Moses took two animals of each kind on the ark*) often goes unnoticed due to strong world-knowledge associations between materials later in the sentence (Erickson & Mattson, 1981), it can be easily detected by clefting the subject *Moses* and placing it under focus (*It was Moses who put two of each kind of animal on the ark*). Our study complements previous findings in this literature by demonstrating that the processing of focus-related semantic meaning signaled by pitch accents can also modulate the ability to detect incongruencies (except in our case, the depth of processing conspired to strengthen misinterpretations). In sum, the interference of the initial misanalysis lingering in memory, as opposed to a lingering local structure due to the failure of syntactic reanalysis, appears to drive misinterpretations in garden path sentences.

---

<sup>17</sup> One mechanistic explanation for this phenomenon could be that the recovery of the correct syntactic structure is what controls the timing of the parser's exit from the reanalysis process. Under this interpretation, if the semantic processor fails to discard the initial misinterpretation within this time frame due to it being strongly encoded in memory, then the misinterpretation persists and interferes with comprehension.

## 5. Conclusion

In this study, we examined the depth of semantic processing as a factor that influences the effects of semantic persistence, and that is independent of previously reported local coherence effects. Depth of semantic processing was examined through known effects of pitch accent on the processing of focus-related semantic meaning. Since focus-marking pitch accents are interpreted with respect to the syntactic position of the accented word, we reasoned that a pitch accent on a temporarily syntactically ambiguous word could lead to an interpretation of focus-related semantic meaning in the incorrect local parse. Specifically, in garden path sentences with early/late-closure ambiguity, we predicted that a pitch accent on the adjunct clause verb would be interpreted as marking contrastive focus in the incorrect late-closure parse, facilitating deeper semantic processing for the associated erroneous interpretation and making it more likely to persist in memory, ultimately resulting in lower accuracy on comprehension questions.

In a speeded auditory comprehension experiment, we found that a nuclear pitch accent on the verb indeed lowers accuracy compared to the baseline condition with a nuclear pitch accent on the adjunct subject, which is positioned outside the region of ambiguity. This effect is significant after controlling for previously reported local coherence effects from transitivity bias and semantic fit, consistent with our prediction that the depth of semantic processing is an independent factor affecting the strength of semantic persistence. A post-hoc analysis of response times suggests against alternative explanations for the effect of pitch accent on accuracy that are based on the pitch accent facilitating or impeding the process of syntactic ambiguity resolution itself, as those accounts would predict a strong corresponding effect of pitch accent on response time. Combined, the data suggest that the pitch accent is immediately interpreted for its focus-related semantic meaning *in the garden path* before the syntactic ambiguity is fully resolved, strengthening the initial erroneous interpretation's resistance to revision in semantic reanalysis.

These findings on the effect of local coherence present a challenge to prior accounts of semantic persistence which claim that the locus of misinterpretations is the failure of syntactic reanalysis, where a sufficiently activated local structure can fail to decay past disambiguation and linger alongside the globally correct structure in the syntactic representation (Ferreira et al., 2004; Tabor et al., 2004). Instead, the results are more consistent with accounts in which syntactic reanalysis may proceed without being accompanied by necessary semantic revisions that would require the semantic processor to discard the initial misinterpretation from memory (Sturt, 2007). Thus, not only is semantic persistence about the failure of reanalysis as opposed to the ill-formedness of the underlying syntactic structure (Slattery et al., 2013), but more specifically, it appears to be about the failure of *semantic* reanalysis (i.e., the failure to sever the semantic commitments to the initial misinterpretation) as opposed to the failure of *syntactic* reanalysis (i.e., the failure to discard the locally coherent structure and prevent it from lingering in the syntactic representation).

These results are interesting not only for sentence processing research, but for prosody research as well. Studies in the real-time processing of pitch accents have traditionally centered on the interpretation of discourse-pragmatic meaning in structurally unambiguous sentences, often in the context of referential ambiguity resolution. Where pitch accents have been studied in relation to syntactic structure building, the focus has been on their effect on the resolution of *global* syntactic ambiguities, such as Carlson and Tyler's (2018) finding that the location of pitch accent biases the height of adverbial-PP attachment when the attachment site is globally ambiguous (see also Schafer et al., 1996; Lee & Watson, 2011). This study contributes to this emerging area of research on pitch accents by examining how they are interpreted when the syntactic position of the accented word is merely *temporarily* ambiguous. Our analysis posits an incremental semantic processor that makes stronger semantic commitments to the erroneous interpretation associated with the initial parse when it is marked for semantic focus via pitch accenting. This is consistent with previous findings for the rapid and immediate processing of focus-semantic meaning from contrastive pitch accents, suggesting that "semantic processing" as talked about in prosody research and sentence processing research may be more interlinked than previously thought. This study serves as yet another testament to the fruitfulness and necessity of synthesizing prosody and sentence processing in psycholinguistics research, as Janet Dean Fodor (2002) emphasized two decades ago.

---

## Data accessibility statement

Stimuli and data from experiments are available at <https://osf.io/u6dq5/>. Analysis scripts and figures for the main auditory comprehension experiment are available at <https://github.com/yjunechoe/Semantic-Persistence>.

## Ethics and consent

All experiments reported in this study were approved by the Northwestern University Institutional Review Board as compliant with the ethical standards for research on human subjects.

## Acknowledgements

The authors would like to thank our two anonymous reviewers, the audience at AMLaP 2020, and members of the Syntax, Semantics and Sentence Processing Lab and the Prosody & Speech Dynamics Lab at Northwestern University for valuable feedback on this project. We especially thank Dan Turner for recording the audio stimuli and Rosemary Dong for assisting in data cleaning for the norming experiments.

## Funding information

This work was supported by the Weinberg College of Arts and Sciences Summary Research Grant and the Academic Year Undergraduate Research Grant from Northwestern University.

## Competing interests

The authors have no competing interests to declare.

## Author contributions

JCh, MY, and JCo conceptualized the experiment and contributed to the writing and analysis. JCh led data collection and scripted the code for the experiments and the analyses.

---

## References

- Abney, S. T. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, *18*, 129–144. DOI: <https://doi.org/10.1007/BF01069051>
- Baker, L., & Wagner, J. L. (1987). Evaluating information for truthfulness: The effects of logical subordination. *Memory & Cognition*, *15*, 247–255. DOI: <https://doi.org/10.3758/BF03197723>
- Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. (2018). Parsimonious mixed models. *ArXiv e-prints*. DOI: <https://doi.org/10.48550/ARXIV.1506.04967>

- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Baumann, S., & Riester, A. (2012). Referential and lexical givenness: Semantic, prosodic and cognitive aspects. In G. Elordieta & P. Prieto (Eds.), *Prosody and meaning* (pp. 119–162). Berlin and New York: Mouton De Gruyter. DOI: <https://doi.org/10.1515/9783110261790.119>
- Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309. DOI: <https://doi.org/10.1017/S095267570000066X>
- Berwick, R. C., & Weinberg, A. (1984). *The grammatical basis of linguistic performance: Language use and acquisition*. Cambridge, MA: MIT Press.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and language development* (pp. 279–352). New York: Wiley & Sons.
- Birner, B. J., & Ward, G. (1998). *Information status and noncanonical word order in English*. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/slcs.40>
- Boersma, P., & Weenink, D. (2020). Praat: Doing phonetics by computer [Computer program]. Version 6.1.06. <http://www.praat.org>
- Büring, D. (2007). Intonation, semantics and information structure. In G. Ramchand & C. Reiss (Eds.), *The Oxford handbook of linguistic interfaces* (pp. 445–474). Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199247455.013.0015>
- Büring, D. (2016). *Intonation and meaning*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199226269.001.0001>
- Calhoun, S. (2006). *Information structure and the prosodic structure of English: A probabilistic relationship*. Edinburgh, United Kingdom: University of Edinburgh dissertation.
- Calhoun, S. (2012). The theme/rheme distinction: Accent type or relative prominence. *Journal of Phonetics*, 40, 329–349. DOI: <https://doi.org/10.1016/j.wocn.2011.12.001>
- Carlson, G. (1984). Thematic roles and their role in semantic interpretation. *Linguistics*, 22(3), 259–280. DOI: <https://doi.org/10.1515/ling.1984.22.3.259>
- Carlson, K., & Tyler, J. C. (2018). Accents, not just prosodic boundaries, influence syntactic attachment. *Language and Speech*, 61(2), 246–276. DOI: <https://doi.org/10.1177/0023830917712282>
- Chodroff, E. & Cole, J. (2019). The phonological and phonetic encoding of information structure in American English nuclear accents. In *Proceedings of the 19th International Congress of Phonetic Sciences*. Melbourne.
- Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures*. Holland: Foris Publications. Reprint. 7th ed. Berlin and New York: Mouton de Gruyter, 1993. DOI: <https://doi.org/10.1515/9783110884166>
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Fernanda Ferreira. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368–407. DOI: <https://doi.org/10.1006/cogp.2001.0752>



- Crocker, M. W. (1996). *Computational psycholinguistics: An interdisciplinary approach to the study of language*. Dordrecht, The Netherlands: Kluwer Academic Publishers. DOI: <https://doi.org/10.1007/978-94-009-1600-5>
- Cutler, A., Dahan, D., & Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40(2), 141–201. DOI: <https://doi.org/10.1177/002383099704000203>
- Erickson, T. D. & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, 20, 540–552. DOI: [https://doi.org/10.1016/S0022-5371\(81\)90165-1](https://doi.org/10.1016/S0022-5371(81)90165-1)
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47, 164–203. DOI: [https://doi.org/10.1016/S0010-0285\(03\)00005-7](https://doi.org/10.1016/S0010-0285(03)00005-7)
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11(1), 11–15. DOI: <https://doi.org/10.1111/1467-8721.00158>
- Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, 30, 3–20. DOI: <https://doi.org/10.1023/A:1005290706460>
- Ferreira, F., Lau, E. F., & Bailey, K. D. G. (2004). Disfluencies, language comprehension, and tree adjoining grammars. *Cognitive Science*, 28, 721–749. DOI: <https://doi.org/10.1016/j.cogsci.2003.10.006>
- Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. *Psychology of Learning and Motivation*, 65, 217–247. DOI: <https://doi.org/10.1016/bs.plm.2016.04.002>
- Ferreira, F., & Patson, N. D. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass*, 1, 71–83. DOI: <https://doi.org/10.1111/j.1749-818X.2007.00007.x>
- Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In *Proceedings of the SPEECH PROSODY 2002 Conference*. Aix-en-Provence, France.
- Fodor, J. D., & Frazier, L. (1980). Is the human sentence parsing mechanism an ATN? *Cognition*, 8(4), 417–459. DOI: [https://doi.org/10.1016/0010-0277\(80\)90003-7](https://doi.org/10.1016/0010-0277(80)90003-7)
- Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, 63(3), 367–386. DOI: <https://doi.org/10.1016/j.jml.2010.06.004>
- Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Storrs, CT: University of Connecticut dissertation.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291–325. DOI: [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1)
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210. DOI: [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)

- Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, 26(5), 505–526. DOI: [https://doi.org/10.1016/0749-596X\(87\)90137-9](https://doi.org/10.1016/0749-596X(87)90137-9)
- Fujita, H. (2021). On the parsing of garden-path sentences. *Language, Cognition and Neuroscience*, 36(10), 1234–1245. DOI: <https://doi.org/10.1080/23273798.2021.1922727>
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37(1), 58–93. DOI: <https://doi.org/10.1006/jmla.1997.2512>
- Gibson, E. (1991). *A computational theory of human linguistic processing: Memory limitations and processing breakdown*. Pittsburgh, PA: Carnegie Mellon University dissertation.
- Gibson, E., & Pearlmutter, N. J. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, 2(7), 262–268. DOI: [https://doi.org/10.1016/S1364-6613\(98\)01187-5](https://doi.org/10.1016/S1364-6613(98)01187-5)
- Gibson, E., & Pearlmutter, N. J. (2000). Distinguishing serial and parallel parsing. *Journal of Psycholinguistic Research*, 29(2), 231–240. DOI: <https://doi.org/10.1023/A:1005153330168>
- Grillo, N., Alexiadou, A., Gehrke, B., Hirsch, N., Paolazzi, C. L., & Santi, A. (2018). Processing unambiguous verbal passives in German. *Journal of Linguistics*, 55, 523–562. DOI: <https://doi.org/10.1017/S0022226718000300>
- Gussenhoven, C. (1992). Sentence accents and argument structure. In I. M. Roca (Ed.), *Thematic structure: Its role in grammar*. Berlin and New York: Foris Publications. DOI: <https://doi.org/10.1515/9783110872613.79>
- Gussenhoven, C. (1999). On the limits of focus projection in English. In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511616983>
- Hare, M., Elman, J. L., Tabaczynski, T., & McRae, K. (2009). The wind chilled the spectators, but the wine just chilled: Sense, structure, and sentence comprehension. *Cognitive Science*, 33, 610–628. DOI: <https://doi.org/10.1111/j.1551-6709.2009.01027.x>
- Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58, 541–573. DOI: <https://doi.org/10.1016/j.jml.2007.06.013>
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, 56(3), 384–409. DOI: <https://doi.org/10.1016/j.jml.2006.09.003>
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2(1), 15–47. DOI: [https://doi.org/10.1016/0010-0277\(72\)90028-5](https://doi.org/10.1016/0010-0277(72)90028-5)
- Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40, 153–194. DOI: <https://doi.org/10.1006/jmla.1998.2620>

- Kurumada, C., Brown, M., Bibyk, S., Pontillo, D. F., & Tanenhaus, M. K. (2014). Is it or isn't it: Listeners make rapid use of prosody to infer speaker meanings. *Cognition*, *133*(2), 335–342. DOI: <https://doi.org/10.1016/j.cognition.2014.05.017>
- Kush, D., & Dillon, B. (2021). Principle B constrains the processing of cataphora: Evidence for syntactic and discourse predictions. *Journal of Memory and Language*, *120*, 104254. DOI: <https://doi.org/10.1016/j.jml.2021.104254>
- Ladd, D. R. (2008). *Intonational phonology, 2nd Edition*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511808814>
- Lau, E. F., & Ferreira, F. (2005). Lingering effects of disfluent material on comprehension of garden path sentences. *Language and Cognitive Processes*, *20*, 633–666. DOI: <https://doi.org/10.1080/01690960444000142>
- Lee, E. K., & Watson, D. G. (2011). Effects of pitch accents in attachment ambiguity resolution. *Language and Cognitive Processes*, *26*(2), 262–297. DOI: <https://doi.org/10.1080/01690965.2010.491650>
- MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*(2), 157–201. DOI: <https://doi.org/10.1080/01690969408402115>
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703. DOI: <https://doi.org/10.1037/0033-295X.101.4.676>
- Marcus, M. P. (1980). *A theory of syntactic recognition for natural language*. Cambridge, MA: MIT Press.
- Marcus, M. P., Hindle, D., & Fleck, M. M. (1983). D-theory: Talking about talking about trees. In *Proceedings of the 21<sup>st</sup> Annual Meeting of the Association for Computational Linguistics*, 129–136. DOI: <https://doi.org/10.3115/981311.981337>
- Paolazzi, C. L., Grillo, N., Alexiadou, A., & Santi, A. (2019). Passives are not hard to interpret but hard to remember: Evidence from online and offline studies. *Language, Cognition and Neuroscience*, *34*(8), 991–1015. DOI: <https://doi.org/10.1080/23273798.2019.1602733>
- Phillips, C. (1996) *Order and structure*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Phillips, C. (2006). The real-time status of island phenomena. *Language*, *82*, 795–823. DOI: <https://doi.org/10.1353/lan.2006.0217>
- Phillips, C., & Lau, E. (2004). Foundational issues. *Journal of Linguistics*, *40*(3), 571–591. DOI: <https://doi.org/10.1017/S0022226704002774>
- Phillips, C., & Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. *Studies in Linguistics*, *6*, 11–47.
- Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 940–961. DOI: <https://doi.org/10.1037//0278-7393.24.4.940>

- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago, IL: The University of Chicago Press.
- Roettger, T. B., Mahrt, T., & Cole, J. (2019). Mapping prosody onto meaning – the case of information structure in American English. *Language, Cognition and Neuroscience*, 34(7), 841–860. DOI: <https://doi.org/10.1080/23273798.2019.1587482>
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75–116. DOI: <https://doi.org/10.1007/BF02342617>
- Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Science*, 6(9), 382–386. DOI: [https://doi.org/10.1016/S1364-6613\(02\)01958-7](https://doi.org/10.1016/S1364-6613(02)01958-7)
- Schafer, A., Carter, J., Clifton, C., & Frazier, L. (1996). Focus in Relative Clause Construal. *Language and Cognitive Processes*, 11(1–2), 135–163. DOI: <https://doi.org/10.1080/016909696387240>
- Schafer, A. J., Speer, S., Warren, P., & White, D. S. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29, 169–182. DOI: <https://doi.org/10.1023/A:1005192911512>
- Schneider, D., & Phillips, C. (2001). Grammatical search and reanalysis. *Journal of Memory and Language*, 45, 308–336. DOI: <https://doi.org/10.1006/jmla.2001.2777>
- Schwarzschild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, 7, 141–177. DOI: <https://doi.org/10.1023/A:1008370902407>
- Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith (Ed.), *The handbook of phonological theory*. London: Blackwell.
- Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, 69, 104–120. DOI: <https://doi.org/10.1016/j.jml.2013.04.001>
- Stowe, L. A. (1986). Evidence for on-line gap-location. *Language and Cognitive Processes*, 1, 227–245. DOI: <https://doi.org/10.1080/01690968608407062>
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, 542–562. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)
- Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition*, 105(2), 477–488. DOI: <https://doi.org/10.1016/j.cognition.2006.10.009>
- Sturt, P., Pickering, M., Scheepers, C., & Crocker, M. (2001). The preservation of structure in language comprehension: Is reanalysis the last resort? *Journal of Memory and Language*, 45, 283–307. DOI: <https://doi.org/10.1006/jmla.2001.2776>
- Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, 50, 355–370. DOI: <https://doi.org/10.1016/j.jml.2004.01.001>

- Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: Digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 431–450. DOI: <https://doi.org/10.1037/0278-7393.30.2.431>
- Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: The MIT Press. DOI: <https://doi.org/10.7551/mitpress/6184.001.0001>
- Traxler, M. J. (2011). Parsing. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 353–364. DOI: <https://doi.org/10.1002/wcs.112>
- Traxler, M. J. (2012). *Introduction to psycholinguistics: Understanding language science*. Oxford: Wiley–Blackwell.
- Traxler, M. J. (2014). Trends in syntactic parsing: Anticipation, Bayesian estimation, and good-enough parsing. *Trends in Cognitive Sciences*, 18(11), 605–611. DOI: <https://doi.org/10.1016/j.tics.2014.08.001>
- Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, 35(3), 454–475. DOI: <https://doi.org/10.1006/jmla.1996.0025>
- Truckenbrodt, H. (1995). *Phonological phrases: Their relation to syntax, focus, and prominence*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Association.
- Wagner, M. (2020). Prosodic Focus. In D. Gutzmann, L. Matthewson, C. Meier, H. Rullmann, & T. E. Zimmermann (Eds.), *The Wiley Blackwell companion to semantics*. Hoboken, NJ: Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781118788516.sem133>
- Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, 25, 905–945. DOI: <https://doi.org/10.1080/01690961003589492>
- Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H\* vs. L+H\*. *Cognitive Science*, 32, 1232–1244. DOI: <https://doi.org/10.1080/03640210802138755>
- Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). DOI: [10.17605/OSF.IO/MD832](https://doi.org/10.17605/OSF.IO/MD832)

